09/869605
JC Rec'd PCT/PTO 02 JUL 2001
4/PRTS

TITLE

A method and a device for providing improved speech intelligibility.

BACKGROUND OF THE INVENTION

The invention relates to a method for the reduction or elimination of an eventual noise contribution to an audio signal, which is picked up by sensor means. The invention further relates to a device for reduction or elimination of an eventual noise contribution to an audio signal, which is picked up by sensor means.

The invention will for practical reasons be explained in connection with a hearing aid. The invention is however not limited to this field as it may be implemented in other technical fields where an audio signal is picked up by sensor means.

Although modern hearing aids have come a long way in providing amplification well suited to the hearing impaired user, many hearing impaired people still have problems understanding speech in noise. The issue of improving the speech intelligibility by improving the ratio between the desired speech signal and the noise has already been addressed in many ways, mostly based on directionality.

As of the present day the commercially available directional devices are either an acoustical two-port design as disclosed in US patent no. 5524056, an electrical combination of the outputs of two omni-directional microphones as disclosed in J. Malsano & W. Hottinger, "A method for electronically beam forming acoustical signals and acoustical sensorapparatus", European Patent EP 0 820 210 A2, or an electrical combination of several microphones into a single highly directional device to be used physically externally to the hearing aid as disclosed in W. Soede, A.J. Berkhout & F.A. Bilsen, 1993, "Development

of a directional hearing instrument based on array technology", J. Acoust. Soc. Am. 94(2), p. 785-798.

These previously known devices are thus characterised by a certain amount of directionality pointing in the forward direction with respect to the user, and characterised by a time-invariant processing of the individual ingoing sensor signals. Although the use of directionality to some extent gives improved speech intelligibility, the directionality is in many situations non-satisfactory and there is a requirement of further improving speech inteligibility.

Directionality is also used in a different group of proposed devices, which make use of adaptive noise cancelling principles as disclosed in B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, R. Hern, J. Zeidler, E. Dong & R. Goodlin, 1975, "Adaptive noise cancelling: Principles and applications", IEEE Proceedings 63, p. 1692-1716. Thus, a number of microphones, which may be mounted either in hearing aids at each ear of the user, on a headband or in a single hearing aid shell, are combined to form estimates of the interfering noise from which the target signal is removed. A signal containing as much target signal as possible is also formed, and these signals are then used as inputs to an adaptive noise canceller. These devices work on the assumption that the target speech signal impinges from a certain predetermined direction, and are characterised by a directionality and a processing of the individual ingoing sensor signals, which is adaptively varying with time. Further, these devices are characterised by the use of the adaptive noise canceller, as disclosed in P.M. Zurek, J.E. Greenberg & P.M. Peterson, 1990, "Adaptive beamforming for noise reduction". United States Patent 4,956,867; P.V.F. Clough & N.A. Lobo, European Patent 0084892; and J. Vanden Berghe & J. Wouters, 1998, "An adaptive noise canceller for hearing aids using two nearby microphones", J. Acoust. Soc. Am. 103(6), p. 3621-3626.

Although the use of adaptive noise cancelling based on directionality to some extend gives improved speech intelligibility, the directionality is in many

situations non-satisfactory in this respect, as it is the case in connection with the use of directionality alone.

Recently, several researchers have published a new principle, which is called independent component analysis (ICA):

1. J. Herault & C. Jutten, "Space or Time adaptive signal processing by neural network models", Neural Networks for Computing, AIP Conference Proceedings 151, Snowbird, Utah, pp. 207-211, 1986.
2. A.J. Bell, United States Patent 5,706,402.
3. A.J. Bell & T.J. Sejnowski, "An information-maximisation approach to blind separation and blind deconvolution", Technical Report no. INC-9501, February 1995, Institute for Neural Computation, UCSD, San Diego, CA 92093-0523.
4. T.W. Lee, M. Girolami, A.J. Bell & T.J. Sejnowski, "A unifying information-theoretic framework for independent component analysis", Computers & Mathematics with applications, 1998 in press.
5. T.W. Lee, A.J. Bell & R. Orglmeister, "Blind source separation of real world signals", Proceedings ICNN, USA, 1997.
6. T.W. Lee, A.J. Bell & R. H. Lambert, "Blind separation of delayed and convolved sources", Advances in Neural Information Processing Systems 9, 1997 MIT Press, Camvridge MA pp. 758-764.
7. P. Smaragdis, 1997, "Efficient blind separation of convolved sound mixtures", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk Mountain House, New Paltz, New York.
8. K. Torkkola, "Blind separation of delayed sources based on information maximation", Proceedings of the IEEE international conference on acoustic, speech and signal processing, May 7-10 1996, GA, USA.
9. K. Torkkola, "Blind separation of convolved ources based on information maximization", IEEE workshop on neural network for signal processing, Kyoto, Japan, Sept 4-6, 1996.
10. J.F. Cardoso, "Equivariant adaptive source separation", IEEE Transactions on Signal Processing 44(12), 1996.

11. S. Amari, A. Cichocki, H. H. Yang, "A new learning algorithm for blind signal separation", Advances in neural information Processing systems 8. MIT Press, 1996.

WO 98/25214 discloses a standard ICA method of signal separation where it is assumed that the sources that generate the sensor signals are stationary. The same assumption applies to the sensor positions and the source signal characteristics. In a real world application these assumptions are obviously violated. WO 98/25214 considers how a standard ICA method should be modified to cope with a non-stationary situation. This comprises a scheme for choosing the learning rate of the ICA algorithm according to online measures of the sensor signal characteristics.

US 5,524,056 discloses aspects where the frequency shaping circuit, which ensures that the frequency responses of directional and omni-directional microphones are similar for frontal incidence of sound. The D-Mic thus contains an arrangement of microphones similar to the two-sensor configuration proposed in our patent, but the combination with an ICA algorithm is not mentioned at all.

DE 195 31 388 C1 discloses similarly to WO 98/25214 an extension of a standard ICA algorithm. In this case an eventual non-linear mixing of the ingoing source signals is taken into account. Sensor arrangement and the question of how to make sure that target signal is chosen as the output of the ICA algorithm, are not discussed at all.

US 5,208,786 discloses only second order statistical moments whereas an ICA algorithm considers higher order statistical moments in the signal separation, The disclosed method is also known as decorrelation or Principal Component Analysis (PCA). Again, sensor arrangement and the question of how to make sure that target speech signal is chosen as the output of the algorithm, are not discussed.

EP 0 565 479 A1 discloses a frequency domain realisation of a standard ICA algorithm. The sensor arrangement and the question of how to make sure that target signal is chosen as the output of the ICA algorithm, are not discussed.

US 5,706,402 discloses what has become standard ICA. Thus it outlines the technical background of the present invention. The sensor arrangement and the question of how to make sure that target signal is chosen as the output of the ICA algorithm are not discussed,

Paper XP-002108812 discloses an experimental application of a standard ICA algorithm in a laboratory setting with speech and noise signals in an anechoic stationary environment. The sensor arrangement is two directional microphones placed on top of each other facing in directions 90 degrees apart. No explanation for the particular choice of sensor configuration is given, and the method with which a particular output signal from the ICA signal separation algorithm has been designated as the target speech signal is not discussed at all.

The objective of ICA is to recover the underlying independent source signals given only sensor observations that are linear mixtures of the original source signals. The only assumption of ICA is that the original source signals are statistically independent, otherwise the statistics of the source signals and the mixing of these into the sensor signals may be unknown. In contrast to correlation-based transformations such as Principal Component Analysis (PCA, I.T. Jolliffe, Principal Component Analysis, 1986, Springer Verlag), which decorrelates signals according to 2^{nd}-order statistics, ICA also reduces higher-order statistical dependencies, in terms of maximising joint output entropy, in order to extract statistically independent signal components.

In the linear blind signal separation problem, N signals, $\mathbf{s}(t) = [s_1(t),...,s_N(t)]^T$, are mixed so that an array of N sensors picks up a set of signals $\mathbf{x}(t) = [x_1(t),...,x_N(t)]^T$, each of which has been mixed, delayed and filtered as follows

$$x_i(t) = \sum_{j=1}^{N} \sum_{k=0}^{M-1} a_{ijk} s_j(t - D_{ij} - k)$$

where D_{ij} are entries in a matrix of delays and $\mathbf{a}_{ij}$ are the M- tap filter coefficients between the jth source and the ith sensor. The problem is to invert this scrambling without knowledge of it, thus recovering the original signals $\mathbf{s}(t)$ given only the $\mathbf{x}(t)$ signals. Finding this inverse scrambling is a challenging task since no informations are provided about the mixing nor the signals (hence the term blind separation). The type of architecture chosen for inverting the scrambling is important and can be made in numerous ways. An accurate architecture for inverting a M-tap filter is an infinite impulse response (IIR) filter with M coefficients. However, IIR filters are limited to have poles inside the unit circle, which imply that a stable filter only exists for a minimum phase system. FIR filters may be used to approximate the inverse solution. Thus the inverse scrambling is performed according to

$$u_i(t) = \sum_{j=1}^{N} \sum_{k=0}^{M-1} w_{ijk} x_j(t - d_{ij} - k)$$

which has filters, $\mathbf{w}_{ij}$, and delays d_{ij}, which supposedly reproduce, at the output $\mathbf{u}(t)$, the original uncorrupted source signals, $\mathbf{s}(t)$, apart from a scaling factor for each signal and a permutation of signals.

Several algorithms have been proposed for the blind separation of linear mixtures. Bell and Sejnowski (3) proposed to learn the separating process by minimising the mutual information between components of $\mathbf{y}(t) = g(\mathbf{u}(t))$, where g is a non-linear function approximating the cumulative probability density function of the sources. They showed that for positively kurtotic

signals (like speech) minimising the mutual information between components of $\mathbf{y}(t)$ is equal to maximising the entropy of $\mathbf{y}(t)$, which can be written as $H(\mathbf{y}) = -E[\ln(f_y(\mathbf{y}))]$, where $f_y(\mathbf{y})$ denotes the probability density function of $\mathbf{y}(t)$. Denoting the determinant of the Jacobian of the whole unmixing process by $|J|$, $f_y(\mathbf{y})$ can be written as $f_x(\mathbf{x})/|J|$ (the Jacobian is a matrix with entries of $\partial y_i / \partial x_j$). Maximising the entropy of the output leads to maximising $E[\ln(|J|)]$, which in turn can be developed into a stochastic gradient ascent rule using instances of $\mathbf{x}(t)$ and $\mathbf{y}(t)$, instead of using the expectation. Thus

$$\Delta \mathbf{W} \propto (1 - 2\mathbf{y}(t))\mathbf{x}(t)^T + [\mathbf{W}^T]^{-1}$$

where $g(u) = 1/(1+e^{-u})$ is used to approximate the cdf.

The algorithm can be made more efficient and independent of the conditioning of the mixing process (matrix) by using the so-called natural gradient instead of the absolute gradient, see Amari (11).
One particular proposed application of ICA is within electroencephalographic (EEG) recording of scalp potentials in humans and related brain activity measurements.

The application of the ICA to hearing aids has also been mentioned in H. Sahlin, "Blind signal separation by second order statistics", Ph.D. thesis, Chalmers University of Technology, Technical Report no. 345, Sweden, 1998, but it has so far never been suggested how the implementation of this technique could be realised in connection with audio systems in order to improve speech intelligibility.

Based on this prior art the objective of the present invention is to provide a method and a device for reducing noise in an audio signal comprising both noise and target signal, which method and device has an increased

functionality and reliability compared with the prior art within the audio field of technology.

SUMMARY OF THE INVENTION

This is achieved by means of a method comprising the steps of:
providing at least two input signals, each having different contents of target signal and noise;
processing the two input signals;
the processing comprising use of an independent component analysis or a similar technique based on the differences of the at least two input signals, hereby determining whether statistical dependent signal elements are present and removing at least part of unwanted signal elements;
outputting a part of the audio signal.

The target signal is defined as the signal coming from in front of the device performing the method, whereas other signals are considered as noise signals. In case of no specific well-defined source signal and cases laying outside the assumed set-up, the noise elimination will disappear meaning that the signal processing strategy will pass the input signals unaltered to the outputs.

In a preferred embodiment the method comprises at least two input signals, which are picked up at least at two mutually distanced locations. Hereby one signal contains the target signal with a higher signal-to-noise ratio than the other input signal. In this case the sensor may be identical.

In a further preferred embodiment the method comprises at least two input signals, which are based on differences in the directionality, preferably a directional and an omni-directional sensor. Hereby it becomes possible to determine the origin of the respective signal components based on the differences of the directionality applied when sensing the two input signals.

The different embodiments show that the differences in signal-to-noise ratio is based on the actual situation of use, which means that it is the mutual differences in signal-to-noise ratio in relation to a desired target signal, which is relevant, and not the signal-to-noise ratio of the sensor means itself.

In a further preferred embodiment two or more output signals are produced and where a possibility exists for switching between the two or more output signals or combinations of these. Alternatively an automatic switching between the two or more output signals according to a predetermined scheme is provided. Hereby it becomes possible to make a choice of the actual influence of the separation method on the noise canceling in the actual situation or at least have a change according to the actual choice of the predetermined scheme, which may be adapted for switching at different listening environments.

The objective of the invention is further achieved by means of a device comprising:

at least two input channels;

signal processing means in connection with the input channels, which provides input signals to the signal processing means;

a receiver in connection with the signal processing means;

the signal processing means being adapted to process the signals by means of an independent component analysis method or a similar method based on differences of signal-to-noise ratios of the input signals in relation to a desired target signal, the processing comprising determining whether statistical dependent signal elements are present and removing at least part of the unwanted signal elements, thereby enhancing other parts of the audio signal.

In a preferred embodiment the device comprises two sensors, preferably microphones, having different signal-to-noise ratio (S/N-ratio). One of the sensors is chosen as the target sensor. Hereby it becomes possible to separate

the respective signal components based on the noise content of the further sensor.

In a preferred embodiment the device comprises a directional microphone and an omni-directional microphone. Hereby the directional microphone will preferably contain the desired output signal based on the position of the user facing the desired audio signal source.

In a preferred embodiment the two microphones are mutually distanced. Hereby it becomes possible to determine the origin of the respective signal components based on a time delay in reception of the two input signals. This makes it possible to make a beamforming of the input signals hereby possibly adding directionality to at least one of the inputs. Other ways of beamforming may be used in this connection.

In a preferred embodiment two or more output signals are produced and means are provided for switching between the two or more output signals or combinations of these. Alternatively means are provided for automatic switching between the two or more output signals according to a predetermined scheme.

Due to the fact that most hearing impaired have a hearing disorder which makes it even more difficult than for normal hearing persons to separate a target signal from the noise, which is often present in a speech situation, the invention is particularly relevant in connection with the technical field of hearing aids.

The invention therefor further relates to a hearing aid comprising: at least two microphones for audio signal input; signal processing means in connection with the microphones; an amplifier in connection with the signal processing means; a receiver in connection with the amplifier for outputting a signal from the amplifier; the signal processing means being adapted to process the signals

by means of an independent component analysis method or a similar method based on the input from the at least two microphones, the processing comprising determining whether statistical dependent signal elements are present and removing at least part of the unwanted signal elements, thereby enhancing other parts of the audio signal.

The hearing aid according to the invention may further comprise the features set forth above, either separate or in combination.

Other fields of relevant use of the invention may be telecommunication or audio systems. In such systems the input and output may be connected to antennas or similar transmission and receiving means or may comprise microphones as input means as in the case of a hearing aid. Other elements of such systems may be standard elements, as these are not influenced by the signal processing according to the invention.

The invention will be described more detailed with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram showing the principles of the invention;

Fig. 2 is a schematic diagram showing the principles of an implemented version of the invention;

Fig. 3 is a schematic diagram showing the principles of an implemented version of the invention;

Fig. 4 is a schematic diagram showing the implemented version of the invention as shown in fig. 2 further implemented in a hearing aid.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The fundamental principle of the invention is schematically shown in fig. 1. The invention is basically a system, e.g. a hearing aid, with two or more sensors and a calculation unit. The calculation unit carries out the separation of the target and noise signals, by using the independence of the mixed signals according to ICA or a similar method comprising the basics of the ICA. The sensors are arranged so that one is positioned to receive sound primarily from a target direction in front, whereas the others have arbitrary characteristics that do not specifically favour the target direction. Hereby, it is possible to use the technique of independent component analysis to separate the desired signal, which impinges from the target direction, from the disturbing noise signals, which impinge from any other directions.

To illustrate the invention, an example is given of a system implementing signal processing as described above. Fig. 2 schematically shows the signal processing system. The system comprises a directional and an omni-directional microphone, and a digital signal processing unit implementing the signal separation algorithm. Using the directional microphone gives the target direction from in front of the user, whereas the omni-directional microphone gives a signal equally representing all signals around the head of the user.

A particularly important property of the independent component analysis is that it separates convolved and delayed source signals, where each independent source signal is defined as a signal which appears in the same way within each mixing process. Another important characteristic about the independent component analysis is that knowledge about the ratio of the source signals within the mixed signals can be used for classifying the separated signals. If for instance one source signal appears with a significantly better signal to noise ratio in one of the sensor signals, this information can be used to ensure that this source signal always will appear in a fixed output. Within the present invention these two characteristics combined with an appropriate placement of at least two sensors are exploited to eliminate signals not coming from in front of the user of the device.

From fig. 3 an embodiment appears, which comprises the features of the embodiment of fig. 2, but where the signal processor produces two output signals. By means of switching means one of the two output signals may be selected for further processing, e.g. amplification, or for output.

From fig. 4 an embodiment appears as a hearing aid according to the invention. The essential components of the hearing aid comprise two microphones, preferably a directional microphone and an omni-directional microphone, and an A/D converter connected to each of the microphones. The A/D converters are connected to a digital signal processor, which is adapted to perform the ICA method on the incoming signals. The signal from the signal processor is then lead to an amplifier and from this through a D/A converter to a receiver for performing the output of the processed signal. The devices of the figs. is in a usual manner powered by means of usual power sources, such as batteries.